Applied Econometrics 2, Basic stats in R

Евгений Орлов

20.11.2014

### Данные

В качестве входных данных взяты цены закрытия по обыкновенным акциям Сбербанка в период с 01.01.2010 по 07.11.2014 включительно.

Источником данных выступила ИТС QUIK.

sber.df <- read.csv("Сбербанк [Price].txt", header=TRUE)  
tail(sber.df)

## X.TICKER. X.PER. X.DATE. X.TIME. X.OPEN. X.HIGH. X.LOW. X.CLOSE.  
## 2997 SBER [TQBR] Daily 20141030 0 73.70 75.13 73.43 74.98  
## 2998 SBER [TQBR] Daily 20141031 0 75.45 76.84 75.23 76.23  
## 2999 SBER [TQBR] Daily 20141103 0 76.07 76.66 75.61 76.29  
## 3000 SBER [TQBR] Daily 20141105 0 75.80 75.94 75.02 75.65  
## 3001 SBER [TQBR] Daily 20141106 0 75.76 76.88 74.75 75.05  
## 3002 SBER [TQBR] Daily 20141107 0 75.16 77.65 74.41 75.53  
## X.VOL.  
## 2997 18099518  
## 2998 15467572  
## 2999 4989221  
## 3000 10108565  
## 3001 10972235  
## 3002 17931891

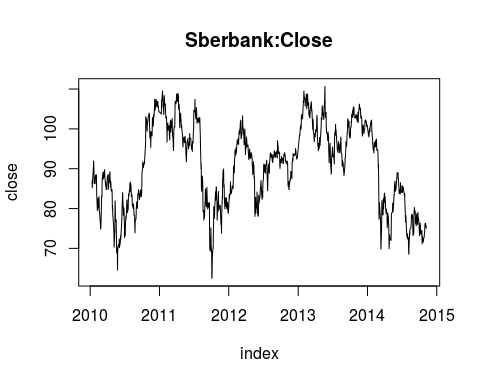
# Цены закрытия начиная с 01.01.2010 по 07.11.2014 включительно  
sber.close <- sber.df[sber.df$X.DATE. >= 20100101, "X.CLOSE."]  
# Длина выборки   
scl <- length(sber.close)  
print(paste0("Размер выборки: ", scl))

## [1] "Размер выборки: 1211"

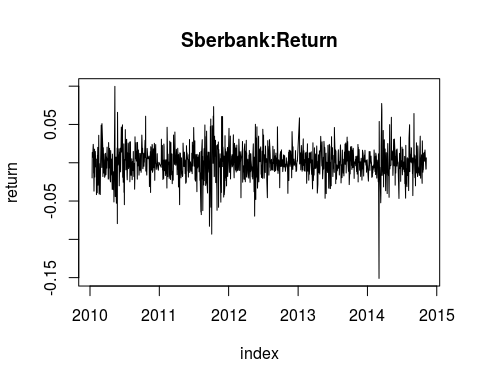
# Вектор доходностей  
sber.ret <- sber.close[2:scl] / sber.close[1:(scl-1)] - 1

Графики цен закрытия и доходностей:

# График цен закрытия  
plot(sber.dates, sber.close,   
 type='l', xlab='index', ylab='close', main='Sberbank:Close')



# График доходностей  
plot(sber.dates[-1], sber.ret,   
 type='l', xlab='index', ylab='return', main='Sberbank:Return')



### Тесты на нормальность

Тестируем распределение доходностей на нормальность с помощью тестов Шапиро-Уилка и Колмогорова-Смирнова:

# Тесты на нормальность  
shapiro.test(sber.ret)

##   
## Shapiro-Wilk normality test  
##   
## data: sber.ret  
## W = 0.9592, p-value < 2.2e-16

ks.test(sber.ret, 'pnorm', mean=mean(sber.ret), sd=sd(sber.ret))

## Warning in ks.test(sber.ret, "pnorm", mean = mean(sber.ret), sd =  
## sd(sber.ret)): ties should not be present for the Kolmogorov-Smirnov test

##   
## One-sample Kolmogorov-Smirnov test  
##   
## data: sber.ret  
## D = 0.0607, p-value = 0.0002704  
## alternative hypothesis: two-sided

Оба теста показывают, что вероятность получить на практике фактическую выборку доходностей в предположении нормальности распределения крайне мала (p-value << 0.01).

Можно сделать вывод, что для достаточно низкого уровня значимости распределение доходностей обыкновенных акций не является нормальным.

При этом, фактические данные не противоречат гипотезе о симметричности распределения (тест Миао, Гэла, Гэствирта(2006)):

# Тест на симметричность распределения  
library(lawstat, quietly=TRUE)

## Loading required package: stats4  
## Loading required package: splines

symmetry.test(sber.ret)

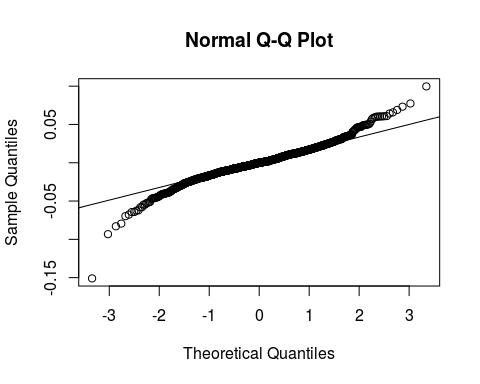
##   
## Test of Symmetry - MGG Test  
##   
## data: sber.ret  
## Test Statistic = 0.236, p-value = 0.8134

Подтвердим результаты, полученные в тестах, визуально.

### Визуальный анализ данных

Построим график "нормальный квантиль-квантиль".

qqnorm(sber.ret)  
qqline(sber.ret)



На графике хорошо заметно, что распределение выборки имеет более толстые хвосты, чем нормальное распреление.

Сопоставим гистограмму выборки с нормальным распределением.

library(fBasics, quietly=TRUE)

##   
## Attaching package: 'fBasics'  
##   
## The following object is masked from 'package:base':  
##   
## norm

histPlot(timeSeries(sber.ret))

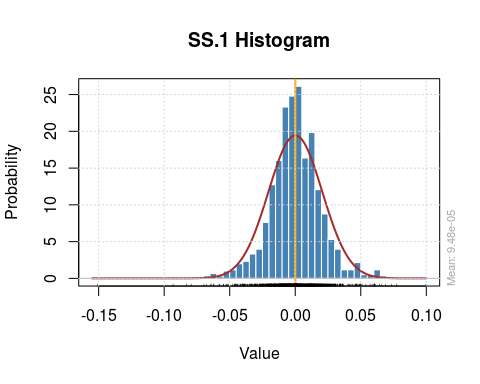


График подчеркивает островершинность распределения (excess kurtosis > 0).

### Вывод

Исходя из полученных результатов, можно сделать вывод, что рассмотренной выборке дневной доходности обыкновенных акций Сбербанка соответстует симметричное островершинное распределение с толстыми (по сравнению с нормальным распределением) хвостами.